Sequentially arranged streptavidin-binding modules as affinity tags

Description

The present invention relates to sequentially arranged streptavidin binding peptide modules which may in particular be used as affinity tags. The affinity tags comprise at least two individual modules capable of mediating avidic binding to streptavidin.

Many projects on resolving genomes of various organisms are close to completion. Thus the human genome has almost completely been sequenced recently and the sequence data has been classified and worked up (published in i) Nature, Vol. 409, 15th February 2001, ii) Science, Vol. 291, No. 5507, 16th February, 2001). One of the next challenges will be the elucidation of the in each case corresponding proteomes, i.e. the elucidation of the protein functions corresponding to the genome and their dynamic interaction for the generation of cellular functions.

With the aid of modern genetic methods it is possible to clone almost any natural gene and to produce the corresponding protein recombinantly in microorganisms or tissue cultures. This allows on the one hand access to proteins which occur in their natural producers only in negligible amounts and which are therefore obtainable from this source only with difficulty, if at all. On the other hand, the recombinant approach also provides a wide variety of possibilities for altering the target protein genetically in order to make it possible thereby to characterize said protein in detail and/or to manipulate it more readily.

A first step in characterizing the target protein is normally its purification from the host proteins, which traditionally can be achieved only by empirical development of a method

specific for each protein. In order to establish a detection method which may be important, for example, for optimizing production processes or else may be helpful for further characterization of the target protein downstream, it has been necessary in the past to prepare a specific antiserum against the desired protein; for this purpose, however, the target protein in turn had to be prepared initially as pure substance. Further methods for downstream analysis of recombinant proteins often require immobilization as irreversible as possible to a solid phase such as, for example, in the wells of a microtiter plate or in the form of ”arrays” on chips (protein chips). Since each protein is a substance with individual properties and therefore each protein is affected in a different way during direct immobilization, it is advantageous for the presentation of proteins in native authentic form if immobilization can be achieved by an independent module in a standardized manner.

A universal solution for these questions is in principle based on the slight modification of a recombinant gene during cloning with nucleotide sequences which code for ”peptide tags”. It is important in this connection that the peptide tag has suitable binding properties for a receptor. In practice, utilization of such a peptide tag is as follows:

After or during expression, the target protein of interest is modified by a peptide tag. The known and well-characterized properties of the peptide tag for binding to its receptor in various assay methods are then available for further analysis of the target protein. Normally, the affinity tag is initially used for purifying the fusion protein by means of an immobilized receptor. It is important for purification by affinity chromatography that the recombinant fusion protein can again be eluted from the solid phase under mild conditions. After purification it may be desirable to be able to utilize the peptide tag for immobilizing the recombinant target protein to a solid phase such as the wall of a

microtiter plate well, for example. Normally, a particularly tight binding is desired here, i.e. the fusion protein should not detach again from the solid phase under any circumstances during the assay method.

A commonly used peptide tag is the His_6 tag. This tag binds to heavy metal ions such as nickel, for example, with chelate formation. A problem occurring when using this tag in the purification of proteins is contamination of the desired protein with the heavy metals used as receptors. Furthermore, chelate formation between heavy metal receptor and His_6 tag represents only low specificity binding. Removing the tag requires high concentrations of imidazole which in many applications may also constitute a problem. All in all, using His_6 tags has so far produced purity degrees of about 80% of the desired protein.

US Patent 5,506,121, for example, has disclosed another class of peptide tags having a specific binding property for streptavidin as receptor. These peptide tags have been denoted Strep-tag and are sold worldwide under this name.

The development of these affinity tags was based on the observation of Devlin et al. (1990) and Lam et al. (1991) that streptavidin is capable of binding peptides at all. The authors regarded the 3 amino acid peptide sequence NH_2-His-Pro-Gln(Met, Asn)-COOH as the smallest motif for streptavidin binding. However, it was impossible to use such peptides alone for practical applications, since the binding affinity was too low (Weber at al., 1992). Only after Schmidt and Skerra (1993) optimized the affinity practical applications did become possible (see also US Patent 5,506,121).

Despite the optimization by Schmidt and Skerra (1993) there were still problems in particular application formats and/or due to different effects of various fusion proteins on streptavidin binding affinity (Schmidt and Skerra, 1994). In

particular, it turned out that the initially most preferred streptavidin-binding peptide having the sequence NH_2-Trp-Arg-His-Pro-Gln-Phe-Gly-Gly-COOH (Strep-tag) could be used only on the free C-terminal end of the recombinant protein fusion partner, since the C-terminal carboxylate group formed an ionic interaction with an arginine residue of streptavidin (Schmidt et al., 1996). For the more general application, the peptide sequence NH_2-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-COOH (Strep-tag II) proved more suitable, since it was possible to utilize this sequence independently of the location on the recombinant fusion protein partner. However, the affinity of the Strep-tag II/streptavidin complex was lower than the affinity of the Strep- tag/streptavidin complex (Schmidt et al., 1996).

Therefore, the receptor streptavidin was optimized with respect to better Strep-tag II binding. It was possible to generate streptavidin muteins which have distinctly higher Strep-tag® II affinity and which have been disclosed in US Patent 6,103,493 and by Voß and Skerra (1997). These streptavidin muteins are commercially available under the name Strep-Tactin. The affinity tag technology based on Strep-tag II/Strep-Tactin interaction (K_d = approx. $1 \cdot 10^{-6}$ M, Voß and Skerra, 1997) is markedly improved for many practical applications compared with the affinity tag technology based on Strep-tag®/streptavidin interaction (K_d = approx. $3.7 \cdot 10^{-5}$ M, Schmidt et al., 1996). Nevertheless, there are still limitations if a particularly tight immobilization is desired. This is the case, for example, for the immobilization of recombinant Strep-tag II fusion proteins on Strep-Tactin-coated microtiter plates or on so-called "protein chips" on which the recombinant proteins, being immobilized in this standardized manner, are then intended to be analyzed in the most complex assay methods or when the recombinant fusion protein which in the extreme problem case is present only in very diluted form and is highly

contaminated with other host proteins is intended to be purified efficiently, in particular in batch format.

It is impossible for one and the same interaction to meet or optimally meet the different demands made on such a peptide tag receptor interaction, i.e. binding as reversible as possible for gentle purification and also binding as irreversible as possible for the immobilization in diagnostic assay systems in microtiter plate format or on protein chips.

Trying, therefore, to solve the problem by generating a generally optimally functioning peptide tag/receptor interaction on monovalent basis presents a dilemma: if binding becomes strong enough for tight binding to surfaces, then it may be impossible to carry out efficient elution under competitive conditions during affinity chromatographic purification. Competitive elution, however, is an elementary condition for being able to carry out affinity chromatographic elution specifically, efficiently and gently (see, for example, Skerra and Schmidt, 1999). The dilemma is caused by each interaction being determined by a binding rate and a dissociation rate. A general principle is that immobilization on surfaces prefers very slow dissociation rates, whereas affinity chromatographic elution by a competitively (competitively means that both ligands can bind separately but not simultaneously) binding agent prefers comparatively rapid dissociation rates. That is to say therefore, an ideal affinity tag should behave under competitive conditions as if it had a fast receptor dissociation rate and should have a very slow dissociation rate under non-competitive conditions.

It was therefore an object of the invention to develop short peptide sequences which can be linked to a recombinant protein without interfering with the function thereof, which make detection using a readily available reagent possible, which display readily controllable binding properties and

which can readily be eluted under competitive conditions despite strong binding affinity to surfaces.

According to the invention, this object is achieved by a comprising at least two streptavidin-binding individual modules or epitopes, wherein the distance between the individual modules is at least 0 and not greater than 50 amino acids and wherein each individual module includes at least the sequence -His-Pro-Baa- where Baa is glutamine, asparagine or methionine.

An achievement of the invention is therefore the use of streptavidin-binding ditags or multitags. This means the sequential arrangement of at least two different or identical streptavidin-binding or/and streptavidin mutein-binding modules (epitopes) which can be employed as fusion partners of the recombinant target protein. Surprisingly, it is possible by such an arrangement to achieve an avidity effect by divalent or multivalent binding of a ditag or multitag to a homotetrameric Strep-Tactin or streptavidin molecule or another streptavidin mutein molecule. Such avidity effects have been known so far primarily for immunoglobulins, since these can adapt through their flexible hinge region to the steric requirements of bivalent binding to two epitopes at the same time. In contrast, the structure of tetrameric streptavidin (Weber et al., 1989) is, compared to antibodies, rather stiff and hardly flexible.

The ditags or multitags of the invention are in particular capable of cooperatively binding to in each case a single streptavidin tetramer or streptavidin dimer. The cooperative binding produces an avidity effect, i.e. increased binding of the peptide tags to a streptavidin receptor. It is assumed that when contacting the peptides of the invention, which comprise at least 2 streptavidin-binding individual modules or epitopes, with a streptavidin receptor, initially an interaction between the individual epitopes and the

streptavidin receptor binding sites takes place in a conventional manner. Formation of the individual bond in this connection is in each case subject to the law of mass action, the strength of the bond being determined by the particular dissociation constants (K_d). However, when breaking a bond between an individual module and a streptavidin receptor, the peptide does not detach from the streptavidin receptor, since the peptide is still bound to the streptavidin receptor by at least one further individual module. Owing to the spatial proximity of the detached individual module to the streptavidin receptor, the detached individual module then rebinds to the streptavidin receptor. Thus, under non-competitive conditions a synergistic or avidic effect can be observed, since rebinding always takes place.

Under competitive conditions, a vacated streptavidin binding site is occupied by another ligand or competitor mostly added in excess so that rebinding cannot take place, owing to displacement effects. This makes it possible to achieve the object of providing 2 different binding affinity strengths at different surrounding conditions (the avid binding strength differing from the monovalent binding strengths more under non-competitive conditions than under competitive conditions).

The peptides of the invention preferably comprise 2, 3 or 4 streptavidin-binding individual modules, particularly preferably 2 or 4 streptavidin-binding individual modules and most preferably 2 streptavidin-binding individual modules.

An essential feature of the peptides of the invention is the fact that they are not 2 separate tags but are a sequential arrangement of at least 2 streptavidin-binding individual modules. In this way, the binding properties are determined by the streptavidin-binding ditag or multitag and are independent of a protein to be fused thereto. In contrast, when using 2 independent tags at the C-terminus and at the N-

terminus of a protein or at the same or different termini of dimeric or multimeric protein complexes, the binding properties are dependent on the particular fusion protein(s) (the domains), in particular on the protein folding thereof.

Using the peptide tags of the invention it is possible i) to generate a markedly stronger binding to immobilized streptavidin or streptavidin muteins by sequential arrangement of at least two streptavidin-binding peptides so that the affinity peptide ii) satisfies the common standards for diagnostic immunological assay methods in microtitre plate format, and the same fusion protein with the same ditag can iii) still be competitively eluted efficiently in the affinity chromatographic purification method. In other words: with respect to its binding properties, in particular binding stability, the ditag behaves in the methods described under i) and ii) like a monotag having distinctly higher binding affinity and in the method described under iii) similar to the individual monotags of which the ditag is composed, when considered separately.

Furthermore, the ditag/streptavidin system is particularly interesting, since there is quite a number of streptavidin-binding peptides but also of streptavidin muteins and since it is thereby possible, by making use of all possible combinations, to generate a particularly wide and yet finely subdivided range of binding activities. If, for example, the sequence NH_2-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-COOH is used, the fusion protein binds particularly strongly to Strep-Tactin®, in any case more strongly than NH_2-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-His-Pro-Gln-Xaa-Xaa-Xaa-COOH which in turn binds more strongly than NH_2-His-Pro-Gln-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-His-Pro-Gln-COOH. The same peptides again bind more poorly to streptavidin. These combination possibilities make it possible to select for any intended application the

combination which produces an affinity for streptavidin or Strep-Tactin® or to another streptavidin mutein, which is ideally suited to the intended use.

The term "streptavidin" as used herein includes wild-type streptavidin, streptavidin muteins and streptavidin-like polypeptides, unless otherwise stated in the individual case. Under wild-type streptavidin (wt-streptavidin), the amino acid sequence disclosed by Argarana et al., Nucleic Acids Res. 14 (1986) 1871-1882 is referred to. Streptavidin muteins are polypeptides which are distinguished from the sequence of wild-type streptavidin by one or more amino acid substitutions, deletions or additions and which retain the binding properties of wt-streptavidin. Streptavidin-like polypeptides and streptavidin muteins are polypeptides which essentially are immunologically equivalent to wild-type streptavidin and are in particular capable of binding biotin, biotin derivative or biotin analogues with the same or different affinity as wt-streptavidin. Streptavidin-like polypeptides or streptavidin muteins may contain amino acids which are not part of wild-type streptavidin or they may include only a part of wild-type streptavidin. Streptavidin-like polypeptides are also polypeptides which are not identical to wild-type streptavidin, since the host does not have the enzymes which are required in order to transform the host-produced polypeptide into the structure of wild-type streptavidin.

The term streptavidin also includes streptavidin tetramers and streptavidin dimers, in particular streptavidin homotetramers, streptavidin homodimers, streptavidin heterotetramers and strepavidin heterodimers. Each subunit normally has a binding site for biotin or biotin analogues or for streptavidin-binding peptides.

Examples of streptavidins or streptavidin muteins are mentioned, for example, in WO 86/02077, DE 19641876 A1, US 6,022,951, WO 98/40396 and WO 96/24606.

Another application to be preferred for a ditag is the particularly efficient purification of recombinant fusion proteins from diluted solutions in batch format (in contrast to column chromatography format). Due to the significantly increased apparent affinity in comparison with the monotag, the protein in the diluted solution becomes more concentrated on the immobile phase and thus less recombinant fusion protein is lost in the sequential washing steps during which the equilibrium always re-establishes itself. As soon as the competitively binding agent is then added in excess, the ditag-carrying recombinant protein can be eluted efficiently in batch format, too.

"Batch format" means in particular those assay formats in which the eluent does not migrate through a column or bed but in which receptors are attached to a solid phase and during each washing step virtually the entire liquid phase is removed. Examples of a batch format are magnetic beads which carry receptors on their surface and are contacted in a washing step in each case with liquid, and this liquid can then be removed again completely or virtually completely in each washing step. Another example of batch format are protein chips or receptors introduced onto microtitre plates or similar plates.

In accordance with the invention, a minimum binding ditag comprises an isolated peptide which is composed of at least 2 individual modules (epitopes), wherein the distance between both modules is at least 0 and not greater than 50 amino acids and wherein each individual module includes at least the sequence -His-Pro-Baa- where Baa is glutamine, asparagine or methionine.

Using the ditags or multitags of the invention makes it possible to achieve higher affinities for the particular streptavidin or streptavidin mutein than using an individual monotag (for example the better binding one if two different tags were used), even if the better binding module of the two individual modules is used as monotag. In addition, it is also possible to generate finer grading of the affinity range than is possible by monotags. Nevertheless it is possible to elute competitively ditag- or multitag-carrying recombinant fusion proteins efficiently from solid phases coated with a receptor. High flexibility in the grading of the affinity range of the tags of the invention can be achieved in a simple manner and can be obtained in particular by the choice of peptide tag (which can be composed of a number of different or/and identical individual modules), by the choice of receptor/streptavidin and by the choice of the competitor for elution under competitive conditions.

Furthermore it is possible to prepare stable dimeric recombinant proteins in the following arrangement: recombinant protein-ditag-streptavidin (mutein) with a total of four binding sites-ditag-recombinant protein. It is also possible to link two different recombinant proteins stably via such an arrangement.

A preferred application of ditag fusion proteins is a) stable binding of the fusion partner to surfaces coated with streptavidin (mutein) and/or b) efficient purification of ditag fusion proteins from diluted solutions, in particular in batch format (in contrast to column chromatography format). Especially the highly parallel purification on a small scale but with high yields and high degrees of purity and from complex mixtures is a great challenge for affinity tag systems in high throughput formats, which has been met by the ditag approach for streptavidin-binding affinity tags.

In the peptide of the invention the two modules mediating binding to streptavidin are at a distance of at least 0 and not more than 50, preferably at a distance of at least 4, more preferably at least 8 and up to preferably not more than 30, more preferably not more than 20 amino acids. Particular preference is given to the distance between the two individual modules mediating binding to streptavidin being 8 or 12 amino acids. The length of the binding modules is preferably at least 3, more preferably at least 4 and most preferably at least 6 and preferably not more than 15, more preferably not more than 12 and most preferably not more than 8 amino acids.

The amino acids located between the individual modules may be any amino acids. They are preferably naturally occurring amino acids but chemically modified amino acids may also be present. Such chemically modified amino acids may in particular be incorporated in the case of an in vitro expression system.

In the peptide of the invention the streptavidin-binding individual modules are present sequentially, i.e. no protein with a biological function is located between the individual modules but, where appropriate, only a certain number of linker amino acids. Preferred linker amino acids are Gly and Ser, in particular chains which exclusively or mainly, for example > 60%, contain Gly and Ser. The linker length can be adjusted to the distance of the binding centers in the particular streptavidin receptor. In a tetrameric streptavidin, for example, two binding sites are located at the front and two binding sites at the back. Preference is therefore given to peptide tags which have 2 or 4 (2 + 2) binding sites and in between linkers which are just suited to bridging the distance of the binding centers. It is, however, also possible to use a tag comprising an uneven number of individual modules such as 3 or 5. Particular preference is given to peptides having the following structure:

Where appropriate, no or 1 to 50 linker amino acid(s) – binding module with 3 to 15 amino acids, in particular 3 to 8 amino acids – linker region with 0 to 20, in particular 8 to 12 linker amino acids – binding module with 3 to 15, in particular 3 to 8 amino acids – where appropriate, another non-functional peptide region with no or 1 to 50 linker amino acid(s) or, where appropriate, no or 1 to 50 linker amino acids – binding module with 3 to 15 amino acids, in particular 3 to 8 amino acids – linker region with 0 to 20, in particular 8 to 12 linker amino acids – binding module with 3 to 15, in particular 3 to 8 amino acids – linker region with 15 to 40, in particular 18 to 25 linker amino acids – binding module with 3 to 15, preferably 3 to 8 amino acids – linker section with 0 to 20, in particular 8 to 12 linker amino acids – binding module with 3 to 15, in particular 3 to 8 amino acids – where appropriate, linker section with no or 1 to 50 linker amino acid(s).

Besides the two individual modules and the amino acids which are located between the individual modules the peptide of the invention may contain further amino acids which adjoin to one side of at least one of the individual modules. The total length of the isolated peptide of the invention is preferably at least 6 amino acids, more preferably at least 20 amino acids and may be preferably up to 500 amino acids, more preferably up to 100 amino acids and most preferably up to 56 amino acids and in particular up to 40 amino acids.

At least one of the two individual modules which mediate binding to streptavidin is preferably selected from one of the sequences: -His-Pro-, -His-Pro-Gln-, -His-Pro-Gln-Phe-, -Oaa-Xaa-His-Pro-Gln-Phe-Yaa-Zaa-, where Oaa is Trp, Lys or Arg, Xaa is any amino acid and preferably a naturally occurring amino acid and where either Yaa and Zaa are both Gly or Yaa is Glu and Zaa is Lys or Arg,

-Trp-Xaa-His-Pro-Gln-Phe-Yaa-Zaa, where Xaa is any amino acid and where either Yaa and Zaa are both Gly or Yaa is Glu and Zaa is Lys or Arg or
-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys.

The preferred sequentially arranged binding modules provide an increased affinity for streptavidin but can still be removed sufficiently in the competitive assay format.

The peptide of the invention particularly preferably has the sequence Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-$(Xaa)_n$-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys, where Xaa is any amino acid and n is an integer from 5 to 20, in particular from 8 to 12, and the sequence Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-$(GlyGlyGlySer)_n$-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys, where n is an integer from 1 to 5 and preferably 2 or 3.

Within the scope of the invention it was found that the peptide sequences of the invention have high binding affinity for streptavidin or nuclear streptavidin (a proteolytic cleavage product of streptavidin) (Bayer, E.A., et al., Biochem. J. 259 (1989), 369-376) and for streptavidin muteins, which affinity is in particular higher than the binding affinities of the individual binding modules; at the same time the said peptides can readily be eluted under competitive conditions.

The individual binding modules of the peptide of the invention preferably have a binding affinity K_d for the particular streptavidin receptor of not more than 10^{-2}, more preferably not more than 10^{-3}, even more preferably not more than 10^{-5} and at least 10^{-13}, more preferably at least 10^{-10}, even more preferably at least 10^{-8} and most preferably at least 10^{-6} M. Elution under competitive conditions is then preferably carried out using a competitor which has a higher affinity for the particular streptavidin receptor, preferably an affinity which is at least one order of magnitude, more

preferably at least two orders of magnitude and most preferably at least three orders of magnitude greater. The binding affinity of streptavidin/biotin, for example, is 4×10^{-14} M. Owing to the sequential arrangement of the inventive ditags and of the avidity effect connected therewith the use of 2 binding modules which in each case have a binding affinity of 10^{-6} M or higher, in particular 10^{-5} M or higher, is in particular possible, and despite the low binding affinities of the individual binding modules a strong binding to the streptavidin receptor is obtained under non-competitive conditions.

Particular preference is given to those streptavidin muteins which are described in US Patent 6,103,493 and also in DE 196 41 876.3. These streptavidin muteins have at least one mutation within the region of amino acid positions 44 to 53, based on the amino acid sequence of wild-type streptavidin. Preference is given to muteins of a minimal streptavidin, which start N-terminally in the region of amino acids 10 to 16 of wild-type streptavidin and end C-terminally in the region of amino acids 133 to 142 of wild-type streptavidin. Examples of such streptavidin muteins have a hydrophobic aliphatic amino acid instead of Glu at position 44, any amino acid at position 45, a hydrophobic aliphatic amino acid at position 46 or/and a basic amino acid instead of Val at position 47. Particular preference is given to streptavidin muteins having the sequence Ile-Gly-Ala-Arg or Val-Thr-Ala-Arg at amino acid positions 44 to 47.

The isolated peptide of the invention is preferably used as label (tag) or affinity tag. The invention therefore further relates to a fusion protein comprising an inventive peptide as described above which has at least two individual modules binding to streptavidin or streptavidin muteins linked to a protein. If the peptide sequence of the invention is present in a fusion protein, this fusion protein, too, has a high

affinity for streptavidin and can at the same time be readily eluted under competitive conditions.

Besides eluting under competitive conditions, i.e. in the presence of another streptavidin ligand, it is also possible to break the receptor:affinity tag interaction by changing the pH (pH shift), which makes simple elution possible. Under acidic conditions at least one histidine residue of the peptide of the invention is protonated resulting in the breaking up of the receptor:affinity tag interaction.

The peptide sequence of the invention may be located at the carboxy terminal end, at the amino terminal end or within the amino acid sequence of the protein, as long as this is not connected with adverse properties, such as, for example, inhibition or destruction of the biological activity if it is desired to maintain said activity.

The protein present in the fusion protein may be both a complete protein and a protein mutant such as, for example, a deletion mutant or substitution mutant, or else may be only part of a protein. The peptide of the invention may be linked to the desired protein directly or via linker or spacer sequences.

The invention further relates to an expression vector which contains a nucleic acid sequence, in particular a DNA sequence, coding for a peptide of the invention and which has one or more restriction cleavage sites 5’ or/and 3’ from this nucleic acid sequence, which allow introduction of another nucleic acid sequence, in particular a DNA sequence, coding for the protein to be expressed or for a protein part. The nucleic acid sequence is preferably under the control of a suitable promoter and, where appropriate, an operator. Preferably, at least one restriction cleavage site immediately adjoins the nucleic acid sequence coding for the peptide of the invention. However, it is also possible to

provide the cleavage sites at some distance so that a spacer or linker region is provided between the peptide of the invention and the protein to be fused thereto. The linker region may also include or represent a cleavage site for a sequence-specific protease such as, for example, enterokinase or factor Xa so that the affinity tag can be cleaved off the desired protein after expression and, where appropriate, purification of the fusion protein.

With the aid of the expression vector of the invention it is made possible to arrange the nucleic acid sequence for a protein of interest with a nucleic acid sequence for the peptide of the invention in a simple manner and to obtain, after expression, a fusion protein of the invention. For example, inserting into a restriction cleavage site 5' from the nucleic acid sequence for the peptide of the invention the nucleic acid sequence for the protein to be fused thereto produces a fusion protein which has the streptavidin affinity-mediating ditag or multitag peptide of the invention at the carboxy terminus. Correspondingly, the affinity tag of the invention is located at the amino terminus if the nucleic acid sequence of the peptide to be fused is inserted into a restriction cleavage site in 3' direction.

The restriction cleavage site in the expression vector of the invention need not necessarily be located directly beside the first or last base of the nucleic acid sequence coding for the peptide. Preferably, however, it ought to be located such that the reading frame is not adversely affected during translation and a linkage of only a few, preferably not more than 10, additional amino acids is formed between the peptide and the amino acid sequence of the protein.

The invention further relates to a method for preparing a recombinant fusion protein, which involves introducing a nucleic acid sequence coding for the above-mentioned fusion protein into a suitable host cell. The preparation may also

be carried out via in vitro expression, a nucleic acid sequence coding for the fusion protein being introduced into a cell extract or a cell lysate. The fusion protein of the invention can be obtained by expressing the nucleic acid sequence. The presence of the expression product can readily be detected via a conjugate of streptavidin or of a streptavidin mutant and a label or/and a solid phase. Furthermore, it is possible to isolate or purify the desired protein as fusion protein in a simple manner by using the streptavidin-binding properties of the ditag or multitag of the invention, streptavidin affinity chromatography being possible to be used, for example.

The nucleic acid sequence is preferably introduced into a suitable host cell or into cell lysate using an expression vector of the invention, which contains a nucleic acid coding for the desired fusion protein.

The conjugate of streptavidin and label preferably comprises a fluorescent label or/and an enzymatic label, in particular alkaline phosphatase or horseradish peroxidase. However, it is in principle possible to use any label which allows detection, for example also direct labels such as, for example, gold or latex particles or other labels known to the person skilled in the art.

An essential advantage of the peptides of the invention and the method of the invention for preparing recombinant fusion proteins is the possibility of purifying the expressed fusion protein readily by affinity chromatography via a column with immobilized streptavidin or an immobilized streptavidin mutant or in a batch process. While the ditags or multitags of the invention show high affinity for the receptor, they can nevertheless advantageously be eluted under very mild competitive conditions, for example by adding biotin or biotin-like compounds and in particular by adding 2-iminobiotin, lipoic acid, hydroxyphenylazobenzoic acid

(HABA), dimethylhydroxy-phenylazobenzoic acid (DM-HABA), diaminobiotin or/and desthiobiotin. Thus, the desired fusion protein can be liberated by competitive elution with streptavidin ligands in a simple manner and under mild conditions. Particular preference is given to the elution with biotin on the microscale (of microgram amounts) in batch format.

Because of the very strong interaction between biotin and streptavidin, elution is particularly efficient when using biotin as competitor. When using biotin as competitor, however, a regeneration of the streptavidin receptor is impossible or only possible with difficulty. Therefore the use of biotin is preferred in particular if regeneration is not required, for example on a small scale (microgram amounts of streptavidin).

For carrying out the elution on a larger scale, including the industrial scale (mg to kg amounts of streptavidin), however, preference is given to using a competitor whose binding affinity for the streptavidin receptor is sufficiently high to effect effective elution but which can nevertheless be removed again from the streptavidin receptor in order to provide a regenerable system. On the large scale particular preference is therefore given to using desthiobiotin, iminobiotin, diaminobiotin, lipoic acid, HABA, or/and DM-HABA as competitor. On the large scale, lowering the pH is also an interesting alternative to competitive elution.

It is possible, for example, to use a streptavidin agarose matrix for streptavidin affinity chromatography. Preference is given to using Sepharose polysaccharide (Pharmacia) or Macro Prep (polymethacrylate; Bio Rad) or POROS (polystyrene (OH modified); PE Biosystems) as support material.

The invention further relates to a nucleic acid coding for a ditag peptide or multitag peptide of the invention.

The invention furthermore comprises using streptavidin or/and a streptavidin mutein as receptor for binding a peptide of the invention.

The ditag or multitag of the invention makes rapid and safe detection of fusion proteins which are obtained, for example, as expression products possible. Furthermore, the fusion protein has exactly adjustable and advantageous properties of binding to streptavidin so that a simple purification of the expression product, which can also be carried out on the industrial scale, is made possible.

With the aid of the expression vector of the invention the expression of a fusion protein of the invention is made easier, and such an expression vector can be used universally for all proteins to be expressed. In the fusion protein the peptide of the invention does not interfere with the biological activity of the other part of the protein and therefore need not necessarily be cleaved off before further use. However, if for particular reasons removal of the peptide is desired, then the expression vector of the invention may also be constructed such that it has between the restriction cleavage site for introducing the nucleic acid sequence for the protein and the sequence coding for the peptide a further nucleic acid sequence coding for a specific protease cleavage site. Thus, after expression and, where appropriate, purification or detection of the expression product, cleavage of the peptide sequence can readily be carried out.

The following examples and the attached figure further illustrate the invention.

Figure 1 shows the binding of cytochrome b562, red, with various affinity tags on Strep-Tactin Sepharose, where cytochrome b562-ditag 3 has been applied on the column

denoted ”L”, cytochrome b562-ditag 2 has been applied on the column denoted ”M” and cytochrome b562-monotag has been applied on the column denoted ”R”.

Figure 1.1 shows the binding behavior of the fusion proteins when eluting with 2 ml of a non-competitive buffer. As Figure 1.1 shows, the monotag fusion protein is already washed out of the column under non-competitive conditions, i.e. the band migrates downwards, while the two ditag fusion proteins remain stably immobilized in the upper part of the column.

Figure 1.2 shows the binding behavior of the fusion proteins when eluting with 12 ml of a non-competitive buffer. Whereas here, too, the two ditag variants remain stably immobilized in the upper part of the column, the band of the monotag fusion protein has already broadened very much and has migrated through the gel.

Figure 1.3 shows the binding behavior of the fusion proteins when eluting with 0.5 ml of a competitive buffer containing desthiobiotin.

Figure 1.4 shows the binding behavior of the fusion proteins when eluting with 1.5 ml of a competitive buffer containing desthiobiotin.

As can be seen from Figures 1.3 and 1.4, the competitive displacement rate is nearly identical for all fusion proteins, meaning that under competitive conditions the two ditag fusion proteins are displaced from the column with the same efficiency as the monotag fusion protein.

Examples

Example 1

Three different recombinant proteins with a monotag of the sequence Trp-Ser-His-Pro-Gln-Phe-Glu-Lys, a ditag 1 of the sequence -Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-$(GlyGlyGlySer)_2$-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys and a ditag 2 of the sequence -Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-$(GlyGlyGlySer)_3$-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys were generated so that overall 9 different recombinant fusion proteins were obtained.

For this experiment, two colored proteins (Green Fluorescent Protein (GFP) from A. Victoria, green and cytochrome b562 from E. coli, red) and one enzyme (Alkaline Phosphatase from E. coli) were used in order to be able to detect the presence of the fusion proteins without any further means or after addition of a chromogenic enzyme substrate by means of the color.

The binding properties of all 9 fusion proteins were kinetically studied in a Biacore-apparatus, where streptavidin (having the wild type sequence Glu-Ser-Ala-Val at positions 44 to 47), streptavidin mutein m1 (having the sequence Val-Thr-Ala-Arg at the positions 44 to 47) or streptavidin mutein m2 (having the sequence Ile-Gly-Ala-Arg at positions 44 to 47) were used as receptor.

In order to determine the dissociation kinetics, it was measured once with competitor (e.g. desthiobiotin) and once without competitor.

The presence of the alkaline phosphatase (AP) in a microtiter plate can easily be indirectly detected through staining with an enzyme substrate. It was found that the AP equipped with ditags showed a significantly stronger binding to walls of microtiter plates coated with Strep-Tactin than the AP equipped with a monotag.

Example 2

Using magnetic beads coated with Strep-Tactin a purification of fusion proteins with ditags was carried out in batch format with different dilution steps. Purification of the fusion proteins on small scale was achieved with high yield and high purity.

Example 3

Binding of cytochrome b562 with various affinity tags to Strep-Tactin Sepharose

E. coli cytochrome b562, red, with C-terminal Strep-tag is known to bind particularly poorly to immobilized streptavidin in comparison with other Strep-tag fusion proteins (Schmidt and Skerra, 1994).

In this experiment, binding of cytochrome b562, with various affinity tags at the C terminus (ditag 3; ditag 2; monotag), to 3 columns (L, M, R) with identical Strep-Tactin® Sepharose material was compared:

L (ditag 3): Cytochrome b562-WSHPQFEKGGGSGGGSGGGSWSHPQFEK-COOH
M (ditag 2): Cytochrome b562-WSHPQFEKGGGSGGGSWSHPQFEK -COOH
R (monotag): Cytochrome b562-WSHPQFEK-COOH
(The underlined areas are the streptavidin-binding modules)

Defined purified amounts of the 3 different fusion proteins (700 µg with ditag 3; 800 µg with ditag 2; 950 µg with monotag) were loaded onto a Strep-Tactin® Sepharose column with a bed volume of 2 ml and a biotin binding capacity of approx. 350 nmol per ml. After the protein solution had completely passed into the column, the latter was washed with buffer W (100 mM Tris-HCl pH 8.0; 150 mM NaCl) (non-competitive conditions). Figure 1 and Figure 2 show the

location of the fusion protein after application of 2 ml and 12 ml of buffer W, respectively. It is clearly visible that the cytochrome b562-monotag fusion protein is already washed out of the column under non-competitive conditions and has already been removed completely from the upper part, while the two ditag variants remain comparatively stably immobilized in the upper part of the column. This finding leads to the immediate conclusion that it is possible to purify significantly larger amounts of cytochrome b562-ditag fusion protein on the same amount of identical column material. After washing with 12 ml of buffer W, the column was treated with buffer E (100 mM Tris-HCl pH 8.0; 150 mM NaCl; 2.5 mM desthiobiotin) (competitive conditions). Figure 3 and 4 show the location of the fusion protein after 0.5 and 1.5 ml of buffer E, respectively. The competitive displacement rate is nearly identical for all variants. This means that the two ditag fusion proteins in this chromatography experiment are displaced from the Strep-Tactin Sepharose column under competitive conditions with the same efficiency as the monotag fusion protein.

References

Devlin, J.J., Panganiban, L.C. & Devlin, P.E. (1990), Random peptide libraries: A source of specific protein binding molecules. Science 249, 404-406.

Lam, K.S., Salmon, E.S., Hersh, E.M., Hruby, V.J., Kazmierski, W.M. & Knapp, R.J. (1991). A new type of synthetic peptide library for identifying ligand-binding activity. Nature 354, 82-84.

Schmidt, T.G.M. & Skerra, A. (1993). The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig F_v fragment. Prot. Engineering 6, 109-122.

Schmidt, T.G.M. & Skerra, A. (1994). One-step affinity purification of bacterially produced proteins by means of the 'Strep tag' and immobilized recombinant core streptavidin. J. Chromatogr. A676, 337-345.

Schmidt, T.G.M., Koepke, J., Frank, R. & Skerra, A. (1996). Molecular interaction between the Strep-tag affinity peptide and its cognate target streptavidin. J. Mol. Biol. 255, 753-766.

Skerra, A. & Schmidt, T.G.M. (1999). Applications of a Peptide Ligand for streptavidin: the Strep-tag. Biomolecular Engineering 16, 79-86.

Voss, S. & Skerra, A. (1997) Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification. Protein Eng. 10, 975-982.

Weber, P.C., Pantoliano, M.W. & Thompson L.D. (1992). Crystal structure and ligand-binding studies of a screened peptide complexed with streptavidin. Biochemistry 31, 9350-9354.

Weber, P.C., M.W. Pantoliano and F.R. Salemme (1995). Crystallographic and thermodynamic comparison of structurally diverse molecules binding to streptavidin. Acta Cryst. D51: 590-596.

Claims

1. Isolated peptide, comprising at least two individual modules, wherein the distance between the individual modules is at least 0 and not greater than 50 amino acids and wherein each individual module includes at least the sequence -His-Pro-Baa- where Baa is glutamine, asparagine or methionine.

2. Isolated peptide according to claim 1, wherein at least one individual module comprises the sequence -His-Pro-Gln.

3. Isolated peptide according to claim 1 or 2, wherein at least one individual module comprises the sequence -His-Pro-Gln-Phe.

4. Isolated peptide according to any of claims 1 to 3, wherein at least one individual module includes at least the sequence -Oaa-Xaa-His-Pro-Gln-Phe-Yaa-Zaa- where Oaa is Trp, Lys or Arg, Xaa is any amino acid and where either Yaa and Zaa are both Gly or Yaa is Glu and Zaa is Lys or Arg.

5. Isolated peptide according to any of claims 1 to 4, wherein at least one individual module includes at least the sequence -Trp-Xaa-His-Pro-Gln-Phe-Yaa-Zaa- where Xaa is any amino acid and where either Yaa and Zaa are both Gly or Yaa is Glu and Zaa is Lys or Arg.

6. Isolated peptide according to any of claims 1 to 5, wherein at least one individual module includes at least the sequence -Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-.

7. Isolated peptide according to any of claims 1 to 7, which includes the sequence -Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-

$(Xaa)_n$-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys- where Xaa is any amino acid and n is either 8 or 12.

8. Isolated peptide according to any of claims 1 to 7, which includes the sequence -Trp-Ser-His-Pro-Gln-Phe-Glu-Lys-$(GlyGlyGlySer)_n$-Trp-Ser-His-Pro-Gln-Phe-Glu-Lys- where n is either 2 or 3.

9. Fusion protein comprising a peptide according to any of claims 1 to 8 linked to a protein.

10. Fusion protein according to claim 9, wherein the protein is selected from the group consisting of a complete protein, a protein mutant, in particular a deletion or substitution mutant, and a protein part.

11. Expression vector comprising a nucleic acid sequence which codes for a peptide according to any of claims 1 to 8 and a restriction cleavage site which adjoins said nucleic acid sequence in 5' or/and 3' direction and which allows the introduction of another nucleic acid sequence coding for a protein to be expressed or a protein part.

12. Method for preparing a recombinant fusion protein, wherein a nucleic acid sequence which codes for a fusion protein according to either of claims 9 and 10 is introduced into a suitable host cell or into a cell lysate or into a cell extract.

13. Method according to claim 12, wherein the suitable host cell is transfected with a vector which contains a nucleic acid coding for a fusion protein according to either of claims 9 and 10.

14. Method for detecting or/and obtaining the fusion protein according to claim 16 or 17 in or from a sample, which comprises contacting the sample with a conjugate of

streptavidin and a label or/and with a conjugate of streptavidin a supporting material.

15. Method according to claim 14, wherein a fluorescent label or/and an enzyme label, in particular alkaline phosphatase or horseradish peroxidase is used.

16. Method according to claim 14 for isolating a protein fused to a peptide according to any of claims 1 to 8 from a sample, which comprises subjecting the sample to a streptavidin affinity chromatography to form a complex between the peptide and streptavidin or/and a streptavidin mutein and eluting the protein by contacting the complex with a streptavidin ligand or/and streptavidin mutein ligand and isolating the protein from the sample.

17. Method according to claim 16, wherein the streptavidin ligand used as competitor comprises the amino acid sequence Trp-X-His-Pro-Gln-Phe-Y-Z where X is any amino acid residue and Y and Z are in each case Gly or where Y is Glu and Z is Arg or Lys.

18. Method according to any of claims 17 or 18, wherein the streptavidin ligand used for eluting the fusion protein is biotin or a derivative thereof, in particular 2-iminobiotin, lipoic acid, hydroxyphenylazobenzoic acid, dimethylhydroxyphenylazobenzoic acid, diaminobiotin or/and desthiobiotin.

19. Nucleic acid coding for a peptide according to any of claims 1 to 8.

20. Use of streptavidin or/and a streptavidin mutein as receptor for binding a peptide according to any of claims 1 to 8.

Abstract

The present invention relates to sequentially arranged streptavidin-binding binding modules which may in particular be used as affinity tags. The affinity tags comprise at least two individual modules capable of mediating avidic binding to streptavidin.